

7. SEQUENCE LISTING

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Hook, Magnus  
Patti, Joseph M.  
House-Pompeo, Karen  
Sthanam, Narayana  
Symersky, Jindrich
- (ii) TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS  
AND METHODS OF USE
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Arnold, White & Durkee
  - (B) STREET: P.O. Box 4433
  - (C) CITY: Houston
  - (D) STATE: Texas
  - (E) COUNTRY: U.S.
  - (F) ZIP: 77210
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US Unknown
  - (B) FILING DATE: Concurrently Herewith
  - (C) CLASSIFICATION: Unknown
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 60/017,678
  - (B) FILING DATE: 16-MAY-1996
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Kitchell, Barbara S.
  - (B) REGISTRATION NUMBER: 33,928
  - (C) REFERENCE/DOCKET NUMBER: TAMK:193
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (512) 418-3000
  - (B) TELEFAX: (512) 474-7577

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 441 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATAACATCTG GGAATAAAC	AACGAATGTT ACGGTTCTATA	AAAGTGAAGC GGGAAACAAGT	60
AGTGTTCCT ATTATAAAAC	GGGAGATATG CTACCAGAAG ATACGACACA	TGTACGATGG	120
TTTTAAATA TTAACAATGA	AAAAAGTTAT GTATCGAAAG ATATTACTAT	AAAGGATCAG	180
ATTCAAGGTG GACAGCAGTT	AGATTTAACG ACATTAACAA	TTAATGTGAC AGGTACACAT	240
AGCAATTATT ATAGTGGACA	AAGTGCAATT ACTGATTTG	AAAAAGCCTT TCCAGGTTCT	300
AAAATAACTG TTGATAATAC	GAAGAACACA ATTGATGTAAC	CAATTCCACA AGGCTATGGG	360
TCATATAATA GTTTTCAAT	TAACACAAA ACCAAAATTA	CGAATGAACA GCAAAAGAG	420
TTTGTAAATA ATTCAACAGC	T		441

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 159 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Arg Gly Ser His His His His His	Gly Ser Ile Thr Ser Gly		
1	5	10	15
Asn Lys Ser Thr Asn Val Thr Val His Lys Ser Glu Ala	Gly Thr Ser		
20	25	30	
Ser Val Phe Tyr Tyr Lys Thr Gly Asp Met Leu Pro Glu Asp	Thr Thr		
35	40	45	
His Val Arg Trp Phe Leu Asn Ile Asn Asn Glu Lys Ser Tyr Val Ser			
50	55	60	
Lys Asp Ile Thr Ile Lys Asp Gln Ile Gln Gly Gln Gln Leu Asp			
65	70	75	80
Leu Ser Thr Leu Asn Ile Asn Val Thr Gly Thr His Ser Asn Tyr Tyr			
85	90	95	

Ser Gly Gln Ser Ala Ile Thr Asp Phe Glu Lys Ala Phe Pro Gly Ser  
 100 105 110  
 Lys Ile Thr Val Asp Asn Thr Lys Asn Thr Ile Asp Val Thr Ile Pro  
 115 120 125  
 Gln Gly Tyr Gly Ser Tyr Asn Ser Phe Ser Ile Asn Tyr Lys Thr Lys  
 130 135 140  
 Ile Thr Asn Glu Gln Gln Lys Glu Phe Val Asn Asn Ser Gln Ala  
 145 150 155

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 849 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GACGATAAAA	ATGGAAAAAT	ACAAAATGGT	GACATGATTA	AAGTGGCATG	GCCGACAAGC	60
GGTACAGTAA	AGATAGAGGG	TTATAGTAAA	ACAGTACCAT	TAACTGTTAA	AGGTGAACAG	120
GTGGGTCAAG	CAGTTATTAC	ACCAGACGGT	GCAACAAATTA	CATTCAATGA	TAAAGTAGAA	180
AAATTAAGTG	ATGTTCGGG	ATTTGCAGAA	TTTGAAGTAC	AAGGAAGAAA	TTAACGCAA	240
ACAAATACTT	CAGATGACAA	AGTAGCTACG	ATAACATCTG	GGAATAAAC	AACGAATGTT	300
ACGGTTCATA	AAAGTGAAGC	GGGAACAAGT	AGTGTTCCT	ATTATAAAC	GGGAGATATG	360
CTACCAGAAG	ATACGACACA	TGTACGATGG	TTTTTAAATA	TTAACATGA	AAAAAGTTAT	420
GTATCGAAAG	ATATTACTAT	AAAGGATCAG	ATTCAAGGTG	GACAGCAGTT	AGATTAAAGC	480
ACATTAAACA	TTAATGTGAC	AGGTACACAT	AGCAATTATT	ATAGTGGACA	AAGTGAATT	540
ACTGATTTG	AAAAAGCCTT	TCCAGGTTCT	AAAATAACTG	TTGATAATAC	GAAGAACACA	600
ATTGATGTAA	CAATTCCACA	AGGCTATGGG	TCATATAATA	GTTCATCAAT	TAACTACAAA	660
ACCAAAATTA	CGAATGAACA	GCACAAAGAG	TTTGTAAATA	ATTCACAAGC	TTGGTATCAA	720
GAGCATGGTA	AGGAAGAAGT	GAACGGAAA	TCATTTAAC	ATACTGTGCA	CAATATTAAT	780
GCTAATGCCG	GTATTGAAGG	TACTGTAAAA	GGTGAATTAA	AAGTTTAAA	ACAGGATAAA	840
GATACCAAG						849

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Arg	Gly	Ser	His	His	His	His	His	Gly	Ser	Asp	Asp	Lys	Val	
1				5					10				15		
Ala	Thr	Ile	Thr	Ser	Gly	Asn	Lys	Ser	Thr	Asn	Val	Thr	Val	His	Lys
			20					25				30			
Ser	Glu	Ala	Gly	Thr	Ser	Ser	Val	Phe	Tyr	Tyr	Lys	Thr	Gly	Asp	Met
	35					40					45				
Leu	Pro	Glu	Asp	Thr	Thr	His	Val	Arg	Trp	Phe	Leu	Asn	Ile	Asn	Asn
	50					55				60					
Glu	Lys	Ser	Tyr	Val	Ser	Lys	Asp	Ile	Thr	Ile	Lys	Asp	Gln	Ile	Gln
	65				70				75		80				
Gly	Gly	Gln	Gln	Leu	Asp	Leu	Ser	Thr	Leu	Asn	Ile	Asn	Val	Thr	Gly
				85				90				95			
Thr	His	Ser	Asn	Tyr	Tyr	Ser	Gly	Gln	Ser	Ala	Ile	Thr	Asp	Phe	Glu
			100					105				110			
Lys	Ala	Phe	Pro	Gly	Ser	Lys	Ile	Thr	Val	Asp	Asn	Thr	Lys	Asn	Thr
	115				120					125					
Ile	Asp	Val	Thr	Ile	Pro	Gln	Gly	Tyr	Gly	Ser	Tyr	Asn	Ser	Phe	Ser
	130				135				140						
Ile	Asn	Tyr	Lys	Thr	Lys	Ile	Thr	Asn	Glu	Gln	Gln	Lys	Glu	Phe	Val
	145				150				155			160			
Asn	Asn	Ser	Gln	Ala	Trp	Tyr	Gln	Glu	His	Gly	Lys	Glu	Glu	Val	Asn
			165					170				175			
Gly	Lys	Ser	Phe	Asn	His	Thr	Val	His	Asn	Ile	Asn	Ala	Asn	Ala	Gly
			180					185			190				
Ile	Glu	Gly	Thr	Val	Lys	Gly	Glu	Leu	Lys	Val	Leu	Lys	Gln	Asp	Lys
				195				200				205			
Asp	Thr	Lys													
		210													

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1500 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCACGAGATA TTTCATCAAC GAATGTTACA GATTAACTG TATCACCGTC	TAAGATAGAA 60
GATGGTGGTA AAACGACAGT AAAATGACG TTGACGATA AAAATGGAAA AATACAAAAT	120
GGTGACATGA TTAAAGTGGC ATGGCCGACA AGCGGTACAG TAAAGATAGA GGGTTATAGT	180
AAAACAGTAC CATTAACTGT TAAAGGTGAA CAGGTGGTC AAGCAGTTAT TACACCAGAC	240
GGTGCAACAA TTACATTCAA TGATAAAGTA GAAAAATTAA GTGATGTTTC GGGATTTGCA	300
GAATTTGAAG TACAAGGAAG AAATTTAACG CAAACAAATA CTTCAGATGA CAAAGTAGCT	360
ACGATAACAT CTGGGAATAA ATCAACGAAT GTTACGGTTC ATAAAAGTGA AGCGGGAAACA	420
AGTAGTGTTC TCTATTATAA AACGGGAGAT ATGCTACCAAG AAGATACGAC ACATGTACGA	480
TGGTTTTAA ATATTAACAA TGAAAAAAAGT TATGTATCGA AAGATATTAC TATAAAGGAT	540
CAGATTCAAG GTGGACAGCA GTTAGATTAA AGCACATTAA ACATTAATGT GACAGGTACA	600
CATAGCAATT ATTATAGTGG ACAAAAGTGCA ATTACTGATT TTGAAAAGC CTTTCCAGGT	660
TCTAAAATAA CTGTTGATAA TACGAAGAAC ACAATTGATG TAACAATTCC ACAAGGCTAT	720
GGGTCAATAA ATAGTTTTTC AATTAACATC AAAACAAAAA TTACGAATGA ACAGCAAAAA	780
GAGTTTGTAA ATAATTCAAA AGCTTGGTAT CAAGAGCATG GTAAGGAAGA AGTGAACGGG	840
AAATCATTAA ATCATACTGT GCACAATATT AATGCTAATG CCGGTATTGA AGGTACTGTA	900
AAAGGTGAAT TAAAAGTTT AAAACAGGAT AAAGATACCA AGGCTCCTAT AGCTAATGTA	960
AAATTTAAC TTTCTAAAAA AGATGGATCA GTTGAAAGG ACAATCAAAA AGAAATTGAG	1020
ATTATAACAG ATGCAAACGG TATTGCTAAT ATTAAAGCGT TGCCTAGTGG AGACTATATT	1080
TTAAAAGAAA TAGAGGCGCC ACGACCGTAT ACATTTGATA AGGATAAAGA ATATCCGTTT	1140
ACTATGAAAG ATACAGATAA TCAGGGATAT TTTACGACTA TTGAAAATGC AAAAGCGATA	1200
GAACCAACAA AAGATGTTTC TGCTCAAAAG GTTGGGAAG GCACTCAAAA AGTGAACCA	1260

ACGATTATT TCAAGTTGTA CAAACAAGAT GACAATCAA	ATACAAACACC AGTAGACAAA	1320
GCAGAGATTA AAAAATTAGA AGATGGAACG ACAAAAGTGA	CATGGTCTAA TCTTCCGGAA	1380
AATGACAAAA ATGGCAAGGC TATTAAATAT TTAGTTAAAG AAGTAAATGC	TCAAGGTGAA	1440
GATACAACAC CAGAAGGATA TACTAAAAAA GAAAATGGTT	TAGTGGTTAC TAATACTGAA	1500

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Arg Gly Ser His His His His His Gly Ser Ala Arg Asp Ile			
1	5	10	15
Ser Ser Thr Asn Val Thr Asp Leu Thr Val Ser Pro Ser Lys Ile Glu			
20	25	30	
Asp Gly Gly Lys Thr Thr Val Lys Met Thr Phe Asp Asp Lys Asn Gly			
35	40	45	
Lys Ile Gln Asn Gly Asp Met Ile Lys Val Ala Trp Pro Thr Ser Gly			
50	55	60	
Thr Val Lys Ile Glu Gly Tyr Ser Lys Thr Val Pro Leu Thr Val Lys			
65	70	75	80
Gly Glu Gln Val Gly Gln Ala Val Ile Thr Pro Asp Gly Ala Thr Ile			
85	90	95	
Thr Phe Asn Asp Lys Val Glu Lys Leu Ser Asp Val Ser Gly Phe Ala			
100	105	110	
Glu Phe Glu Val Gln Gly Arg Asn Leu Thr Gln Thr Asn Thr Ser Asp			
115	120	125	
Asp Lys Val Ala Thr Ile Thr Ser Gly Asn Lys Ser Thr Asn Val Thr			
130	135	140	
Val His Lys Ser Glu Ala Gly Thr Ser Ser Val Phe Tyr Tyr Lys Thr			
145	150	155	160
Gly Asp Met Leu Pro Glu Asp Thr Thr His Val Arg Trp Phe Leu Asn			
165	170	175	
Ile Asn Asn Glu Lys Ser Tyr Val Ser Lys Asp Ile Thr Ile Lys Asp			

180

185

190

Gln Ile Gln Gly Gly Gln Gln Leu Asp Leu Ser Thr Leu Asn Ile Asn  
 195 200 205

Val Thr Gly Thr His Ser Asn Tyr Tyr Ser Gly Gln Ser Ala Ile Thr  
 210 215 220

Asp Phe Glu Lys Ala Phe Pro Gly Ser Lys Ile Thr Val Asp Asn Thr  
 225 230 235 240

Lys Asn Thr Ile Asp Val Thr Ile Pro Gln Gly Tyr Gly Ser Tyr Asn  
 245 250 255

Ser Phe Ser Ile Asn Tyr Lys Thr Lys Ile Thr Asn Glu Gln Gln Lys  
 260 265 270

Glu Phe Val Asn Asn Ser Gln Ala Trp Tyr Gln Glu His Gly Lys Glu  
 275 280 285

Glu Val Asn Gly Lys Ser Phe Asn His Thr Val His Asn Ile Asn Ala  
 290 295 300

Asn Ala Gly Ile Glu Gly Thr Val Lys Gly Glu Leu Lys Val Leu Lys  
 305 310 315 320

Gln Asp Lys Asp Thr Lys Ala Pro Ile Ala Asn Val Lys Phe Lys Leu  
 325 330 335

Ser Lys Lys Asp Gly Ser Val Val Lys Asp Asn Gln Lys Glu Ile Glu  
 340 345 350

Ile Ile Thr Asp Ala Asn Gly Ile Ala Asn Ile Lys Ala Leu Pro Ser  
 355 360 365

Gly Asp Tyr Ile Leu Lys Glu Ile Glu Ala Pro Arg Pro Tyr Thr Phe  
 370 375 380

Asp Lys Asp Lys Glu Tyr Pro Phe Thr Met Lys Asp Thr Asp Asn Gln  
 385 390 395 400

Gly Tyr Phe Thr Thr Ile Glu Asn Ala Lys Ala Ile Glu Lys Thr Lys  
 405 410 415

Asp Val Ser Ala Gln Lys Val Trp Glu Gly Thr Gln Lys Val Lys Pro  
 420 425 430

Thr Ile Tyr Phe Lys Leu Tyr Lys Gln Asp Asp Asn Gln Asn Thr Thr  
 435 440 445

Pro Val Asp Lys Ala Glu Ile Lys Lys Leu Glu Asp Gly Thr Thr Lys  
 450 455 460

Val	Thr	Trp	Ser	Asn	Leu	Pro	Glu	Asn	Asp	Lys	Asn	Gly	Lys	Ala	Ile
465					470					475					480
Lys	Tyr	Leu	Val	Lys	Glu	Val	Asn	Ala	Gln	Gly	Glu	Asp	Thr	Thr	Pro
				485					490						495
Glu	Gly	Tyr	Thr	Lys	Lys	Glu	Asn	Gly	Leu	Val	Val	Thr	Asn	Thr	Glu
				500					505						510

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 345 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met	Arg	Gly	Ser	His	His	His	His	His	Gly	Ser	Met	Val	Ala	Ala
1				5										15

Asp	Ala	Pro	Ala	Ala	Gly	Thr	Asp	Ile	Thr	Asn	Gln	Leu	Thr	Asn	Val
					20			25						30	

Thr	Val	Gly	Ile	Asp	Ser	Gly	Thr	Thr	Val	Tyr	Pro	His	Gln	Ala	Gly
				35				40						45	

Tyr	Val	Lys	Leu	Asn	Tyr	Gly	Phe	Ser	Val	Pro	Asn	Ser	Ala	Val	Lys
				50			55							60	

Gly	Asp	Thr	Phe	Lys	Ile	Thr	Val	Pro	Lys	Glu	Leu	Asn	Leu	Asn	Gly
				65			70			75				80	

Val	Thr	Ser	Thr	Ala	Lys	Val	Pro	Pro	Ile	Met	Ala	Gly	Asp	Gln	Val
					85				90					95	

Leu	Ala	Asn	Gly	Val	Ile	Asp	Ser	Asp	Gly	Asn	Val	Ile	Tyr	Thr	Phe
				100				105					110		

Thr	Asp	Tyr	Val	Asn	Thr	Lys	Asp	Asp	Val	Lys	Ala	Thr	Leu	Thr	Met
					115			120						125	

Pro	Ala	Tyr	Ile	Asp	Pro	Glu	Asn	Val	Lys	Lys	Thr	Gly	Asn	Val	Thr
				130			135						140		

Leu	Ala	Thr	Gly	Ile	Gly	Ser	Thr	Thr	Ala	Asn	Lys	Thr	Val	Leu	Val
				145			150				155			160	

Asp	Tyr	Glu	Lys	Tyr	Gly	Lys	Phe	Tyr	Asn	Leu	Ser	Ile	Lys	Gly	Thr
				165				170					175		

Ile Asp Gln Ile Asp Lys Thr Asn Asn Thr Tyr Arg Gln Thr Ile Tyr  
 180 185 190  
 Val Asn Pro Ser Gly Asp Asn Val Ile Ala Pro Val Leu Thr Gly Asn  
 195 200 205  
 Leu Lys Pro Asn Thr Asp Ser Asn Ala Leu Ile Asp Gln Gln Asn Thr  
 210 215 220 -  
 Ser Ile Lys Val Tyr Lys Val Asp Asn Ala Ala Asp Leu Ser Glu Ser  
 225 230 235 240  
 Tyr Phe Val Asn Pro Glu Asn Phe Glu Asp Val Thr Asn Ser Val Asn  
 245 250 255  
 Ile Thr Phe Pro Asn Pro Asn Gln Tyr Lys Val Glu Phe Asn Thr Pro  
 260 265 270  
 Asp Asp Gln Ile Thr Thr Pro Tyr Ile Val Val Val Asn Gly His Ile  
 275 280 285  
 Asp Pro Asn Ser Lys Gly Asp Leu Ala Leu Arg Ser Thr Leu Tyr Gly  
 290 295 300  
 Tyr Asn Ser Asn Ile Ile Trp Arg Ser Met Ser Trp Asp Asn Glu Val  
 305 310 315 320  
 Ala Phe Asn Asn Gly Ser Gly Ser Gly Asp Gly Ile Asp Lys Pro Val  
 325 330 335  
 Val Pro Glu Gln Pro Asp Glu Gln Ala  
 340 345

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 139 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Arg Gly Ser His His His His His His Gly Ser Glu Gly Gly Gln  
 1 5 10 15  
 Asn Ser Gly Asn Gln Ser Phe Glu Glu Asp Thr Glu Glu Asp Lys Pro  
 20 25 30  
 Lys Tyr Glu Gln Gly Asn Ile Val Asp Ile Asp Phe Asp Ser Val  
 35 40 45

Pro Gln Ile His Gly Gln Asn Lys Gly Asn Gln Ser Phe Glu Glu Asp  
50 55 60

Thr Glu Lys Asp Lys Pro Lys Tyr Glu His Gly Gly Asn Ile Ile Asp  
65 70 75 80

Ile Asp Phe Asp Ser Val Pro His Ile His Gly Phe Asn Lys His Thr  
85 90 95 -

Glu Ile Ile Glu Glu Asp Thr Asn Lys Asp Lys Pro Ser Tyr Gln Phe  
100 105 110

Gly Gly His Asn Ser Val Asp Phe Glu Glu Asp Thr Leu Pro Lys Val  
115 120 125

Ser Gly Gln Asn Glu Phe Asp Ile Lys Leu Asn  
130 135